Medical records appear to be an area in which authentication is important. Steganographic principles—applied either to film-based records or to the microtopology of documents—can be employed to provide some protection against tampering.

Many industries, e.g. automobile and airline, rely on tags to mark critical parts. Such tags, however, are easily removed, and can often be counterfeited. In applications wherein better security is desired, industrial parts can be steganographically marked to provide an inconspicuous identification/authentication tag.

In various of the applications reviewed in this specification, different messages can be steganographically conveyed by different regions of an image (e.g. different regions of an image can provide different internet URLs, or different regions of a photocollage can identify different photographers). Likewise with other media (e.g. sound).

Some software visionaries look to the day when data blobs will roam the datawaves and interact with other data blobs. In such an era, it will be necessary for such blobs to have robust and incorruptible ways of identifying themselves. Steganographic techniques again hold much promise here.

Finally, message changing codes—recursive systems in which steganographically encoded messages actually change underlying steganographic code patterns—offer new levels of sophistication and security. Such message changing codes are particularly well suited to applications such as plastic cash cards where time-changing elements are important to enhance security.

Again while applicant prefers the particular forms of steganographic encoding detailed above, the diverse applications disclosed in this specification can largely be practiced with other steganographic marking techniques, many of which are known in the prior art. Likewise, while the specification has focused on applications of this technology to images, the principles thereof are generally equally applicable to embedding such information in audio, physical media, or any other carrier of information.

Finally, while the specification has been illustrated with particular embodiments, it will be recognized that elements, components and steps from these embodiments can be recombined in different arrangements to serve different needs and applications, as will be readily apparent to those of ordinary skill in the art.

In view of the wide variety of implementations and applications to which the principles of this technology can be put, it should be apparent that the detailed embodiments are illustrative only and in no way limit the scope of my invention. Instead, I claim as my invention all such embodiments as come within the scope and spirit of the following claims and equivalents thereto.

I claim:

1. A method for surveying distribution of proprietary empirical data sets, such as audio, image, or video data, on computer sites accessible via the internet, comprising:

automatically downloading data, including empirical data sets, from a plurality of computer sites over the internet;

for each of a plurality of empirical data sets obtained by said downloading operation, automatically screening same to identify the potential presence of identification data steganographically encoded therein;

for each of a plurality of empirical data sets screened by said screening operation, discerning identification data, if any, steganographically encoded therein; and

generating a report identifying steganographically encoded empirical data sets identified by the foregoing steps, and the site from which each was downloaded;

wherein there is calibration data steganographically encoded within at least one empirical data set, said calibration data having one or more known properties facilitating identification thereof during the discerning step;

the method including identifying the calibration data within the empirical data set and using data obtained thereby to aid in discerning the identification data from the empirical data set;

wherein the empirical data set has been corrupted since being encoded, said corruption including a process selected from the group consisting of: misregistration and scaling of the empirical data set;

the method further including using said data to compensate for said corruption, wherein the identification data can nonetheless be recovered from the empirical data set notwithstanding said corruption.

2. The method of claim 1 which includes providing a master code signal, and using said code signal in discerning said steganographically encoded identification data from said screened empirical data sets.

3. The method of claim 2 in which said master code signal has the appearance of unpatterned snow if represented in the pixel domain.

4. The method of claim 1 in which said discerning of identification data from said downloaded empirical data is accomplished without previous knowledge of the audio, image, or video information represented thereby.

5. The method of claim 1 which includes identifying proprietors of empirical data sets by reference to identification data steganographically discerned therefrom, and reporting to said proprietors the sites from which their empirical data sets were downloaded.

6. The method of claim 5 in which said identification data includes information in addition to data identifying said proprietor, and the method includes providing said additional data to said proprietors.

7. The method of claim 5 in which said identification data is a serial number index into a registry database containing names and contact information for proprietors identified by said identification data.

8. The method of claim 1 in which the empirical data sets include image data, and the method includes:

converting said image data to pixel form, if not already in said form; and

performing a plurality of statistical analyses on said pixel form image data to discern the identification data therefrom.

9. The method of claim 8 in which each statistical analysis includes analyzing a collection of spaced apart pixels to decode a single, first bit of the identification data therefrom, said analysis to decode the first bit encompassing not just said spaced apart pixels, but also pixels adjacent thereto, said adjacent pixels not being encoded with said first bit.

10. A method for surveying distribution of proprietary empirical data sets on computer sites accessible via the internet, comprising:

providing a master code signal useful for detecting steganographic coding within empirical data sets;

automatically downloading data, including empirical data sets, from a plurality of computer sites over the internet;

for each of a plurality of empirical data sets obtained by said downloading operation, discerning certain identification data, if any, steganographically encoded therein, said discerning employing said master code signal as a decoding key; and generating a report identifying steganographically encoded empirical data sets identified by the foregoing steps, and the site from which each was downloaded;

wherein there is calibration data steganographically encoded within at least one empirical data set, said calibration data having one or more known properties facilitating identification thereof during the discerning step;

the method including identifying the calibration data within the empirical data set and using data obtained thereby to aid in discerning the identification data from the empirical data set;

wherein the empirical data set has been corrupted since being encoded, said corruption including a process selected from the group consisting of: misregistration and scaling of the empirical data set;

the method further including using said data to compensate for said corruption, wherein the identification data can nonetheless be recovered from the empirical data set notwithstanding said corruption.

11. The method of claim 10 which includes automatically screening each of a plurality of said empirical data sets obtained by said downloading operation, to identify the potential presence of identification data steganographically encoded therein and, for those data sets that pass said screening process, discerning identification data, if any, steganographically encoded therein.

12. The method of claim 10 in which said master code signal has the appearance of unpatterned snow if represented in the pixel domain.

13. The method of claim 10 in which said discerning of identification data from said downloaded empirical data is accomplished without previous knowledge of the audio, image, or video information represented thereby.

14. The method of claim 10 which includes identifying proprietors of empirical data sets by reference to identification data steganographically discerned therefrom, and reporting to said proprietors the sites from which their empirical data sets were downloaded.

15. The method of claim 14 in which said identification data includes information in addition to data identifying said proprietor, and the method includes providing said additional data to said proprietors.

16. The method of claim 14 in which said identification data is a serial number index into a registry database containing names and contact information for proprietors identified by said identification data.

17. The method of claim 10 in which the empirical data sets include image data, and the method includes:

converting said image data to pixel form, if not already in said form; and

performing a plurality of statistical analyses on said pixel form image data to discern the identification data therefrom.

18. The method of claim 17 in which each statistical analysis includes analyzing a collection of spaced apart pixels to decode a single, first bit of the identification data therefrom, said analysis to decode the first bit encompassing not just said spaced apart pixels, but also pixels adjacent thereto, said adjacent pixels not being encoded with said first bit.

* * * * *